Available online at www.sciencedirect.com
SCIENCE DIRECT•
ELSEVIER Computers and Mathematics with Applications 52 (2006) 235–257
An International Journal
computers & mathematics with applications
www.elsevier.com/locate/camwa

Reflections on Optimality and Dynamic Programming

E. A. GALPERIN
Department of Mathematics
Université du Québec à Montréal
C.P. 8888, Succ. Centre Ville
Montréal, Québec H3C 3P8, Canada
galperin.efim@uqam.ca

Abstract—Discrete models and continuous control systems are considered in regard to optimality of their trajectories. Some aspects of the principle of optimality [1, p. 83] are analyzed, and it is shown to imply *total optimality*, that is, the optimality of every part of an optimal trajectory. Certain autonomous systems with free admissible variations possess this property.

Nonautonomous optimal systems are not, in general, totally optimal, in which case the principle of optimality is not valid. A modification is proposed for the derivation of the main functional equation to demonstrate that dynamic programming and its functional equations are valid also in the case of nonoptimal remaining trajectories under a certain *contiguity condition* that is defined and analyzed in the paper.

Control systems with incomplete information or structural limitations on controls do not, in general, satisfy the contiguity condition. Control problems for such systems may have optimal solutions which, however, cannot be obtained by dynamic programming. This fact is shown in an example of a widely used engineering system for which an optimal trajectory has all its remaining parts nonoptimal and noncontiguous to the optimal trajectory. The paper presents theoretical justification of dynamic programming for contiguous systems that do not conform to the principle of optimality. Examples are presented to illustrate the results which open new avenues in modeling and optimization of general (not totally optimal) control systems.

Keywords—Dynamic programming, Total optimality, Contiguity, Control theory.

1. INTRODUCTION

In the ten-line section from [1, p. 83] we read

> "3. The Principle of Optimality
>
> In each process, the functional equation governing the process was obtained by an application of the following intuitive:
>
> *Principle of Optimality. An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*
>
> The mathematical transliteration of this simple principle will yield all the functional equations we shall encounter throughout the remainder of the book. A proof by contradiction is immediate."

0898-1221/06/$ - see front matter Typeset by $\mathcal{AMS}$-TEX
doi:10.1016/j.camwa.2006.08.015

This short statement has laid the foundation of a new direction in operations research, resulting in "a mathematical method called 'dynamic programming'.... In addition to having an impact on applied mathematics and physical sciences, dynamic programming is widely regarded as a major contributor to economic and business management." (From [2].)

In control literature, the mathematical transliteration of the above principle for deterministic processes is related mostly to the following types of functional equations that we reproduce from [1,3], using the simplest representations in authors' notations.

1.1. Discrete Process [3, Part IV, Ch. VIII, Section 8, formula (8.3)]

$$f_{n+1}(p) = \max_q [R(p,q) + f_n(T(p,q))], \qquad n = 0, 1, \dots; \quad f_0(\cdots) = 0, \tag{1.1}$$

where p is the initial state vector for $n = 0$, q is a decision (control) vector corresponding to the first n steps (transitions), $R(p,q)$ is the revenue obtained after the first transition $n = 1$, $T(p,q) = p^1$ is the state vector after one transition, and $f_n(p)$ is the maximum revenue obtained after n steps corresponding to an optimal policy q^o defined by the max-operator in (1.1). Optimal policy q^o need not be unique, however, the optimal value $f_n(p)$ is unique. Clearly, for some problems, e.g., expense minimization, one should write *min* in (1.1), but the structure remains the same. Usually, $q \in S$, $p \in D$, $T(p,q) \in D$, *for all* $q \in S$, *all* $p \in D$, where S and D are bounded closed regions [1, Ch. III, Section 4]; if this is not the case, then *max/min* should be replaced by *sup/inf* [3, Part IV, Ch. VIII, Section 8].

1.2. Continuous Process

In this case, the sequence $f_n(p)$ in (1.1) is replaced by a single cost function $f(p)$, and if $R(p,q)$ is included in $f_n(p)$, the simple functional equation is obtained [1, Ch. III, Section 4, equation (5)]

$$f(p) = \max_q f(T(p,q)). \tag{1.2}$$

1.3. New Formalization of the Calculus of Variations [1, Ch. IX]

For a problem of optimal control with variable final time

$$\max J(u) = \int_0^T F(x,u)\,dt, \qquad \frac{dx}{dt} = G(x,u), \qquad x(0) = c, \tag{1.3}$$

denote

$$\max J(u) = f(c,T), \tag{1.4}$$

and split the integral in (1.3) over two segments $[0,S]$ and $[S,T]$. Then, similarly to (1.1),(1.2) where $R(p,q)$ corresponds to the integral over $[0,S]$, we get according to the principle of optimality

$$f(c,T) = \max_{u[0,T]} \left[\int_0^S F(x,u)\,dt + \int_S^T F(x,u)\,dt \right] \tag{1.5a}$$

$$= \max_{u[0,S]} \left[\int_0^S F(x,u)\,dt + f(c(S), T-S) \right], \tag{1.5b}$$

for all $S \in (0, T)$. For small S, this yields

$$f(c,T) = \max_{u[0,S]} \left[F(x,u)S + f(c,T) + Sf_c \frac{dc(S)}{dS} - Sf_T + o(S) \right]. \tag{1.6}$$

In (1.6) the terms $f(c,T)$ cancel out and, dividing by S, we get, to the first order, as $S \to 0$,

$$0 = \max_{u(0)} [F(c,u(0)) + f_c G(c,u(0)) - f_T]. \tag{1.7}$$

Denoting $u(0) = v$ and in view that f_T, due to (1.4), does not depend on $u(\cdot)$, we can rewrite (1.7) as follows:

$$f_T = \max_v [F(c,v) + f_c G(c,v)]. \tag{1.8}$$

By the principle of optimality, derivation (1.5)–(1.8) can be done starting from any intermediate point $c := x(t) = x$, $0 \le t < T$, so one can replace c, v with x, u in (1.8), yielding

$$f_T(x,T) = \max_u [F(x,u) + f_x(x,T)G(x,u)]. \tag{1.9}$$

Denoting the optimal control by u^o and assuming that u^o is in the interior of the admissible region, one can see that, by classical analysis, the derivative of the bracket in (1.9) with respect to u should be zero at u^o, so that we come to the system, cf. [1, Ch. IX, Section 6],

$$f_T = F(x,u^o) + f_x G(x,u^o), \tag{1.10}$$

$$0 = F_u(x,u^o) + f_x G_u(x,u^o). \tag{1.11}$$

REMARK 1.1. Relations (1.10),(1.11), and (1.17),(1.18) below, are only necessary for max or min in (1.3)–(1.9), and in (1.12)–(1.16) below. They are usually called *stationarity*, or *extremality*, conditions. Their solution $u^o(x)$, or $u^o(x,t)$ for (1.17),(1.18), if it exists, is called *stationary* or *extremal* control (policy), and the corresponding trajectory of (1.3) is called a *stationary* trajectory, or an *extremal* (sometimes, the term extremal is reserved for optimal elements without distinction between max or min). Extremal controls (policies) may exist without being optimal, see Example 6.1 below. To assure optimality of some kind, sufficient conditions should be considered. Usually, they are provided by strict local convexity (concavity) for min (max) of the functional $J(u)$ in a neighborhood of the extremal policy (control), plus the absence of conjugate points (Jacobi condition). Since every optimal policy (trajectory) is also extremal (not vice versa), the principle of optimality naturally translates into a principle of extremality, or stationarity, with the meaning that PDEs (1.10),(1.11) valid for the whole process (trajectory) are valid also for every remaining part thereof.

For fixed $T \le \infty$ (with convergent integral in (1.3) if $T = \infty$), the cost function $f(\cdot)$ in equations (1.4)–(1.9) does not depend on T, so that $f_T = 0$ in (1.6)–(1.10). In this case, system (1.10),(1.11) is homogeneous and, if derivative f_x exists, the determinant in (1.10),(1.11) is zero, yielding the equation that defines the extremal feedback control $u^o(x)$, cf. [1, Ch. IX, Section 3].

If T is variable and one wishes to find an extremal feedback $u^o(x,T)$ depending on the parameter T, then (1.10),(1.11) is solved for f_T, f_x and a partial differential equation for $u^o(x,T)$ is provided by the theorem on mixed partial derivatives: $f_{Tx} = f_{xT}$, cf. [1, Ch. IX, Section 6].

REMARK 1.2. Following [1], we do not specify dimensions $x \in R^n$, $u \in R^m$. If $n > 1$, then (1.6) above and (1.14) below represent the first-order terms of the multidimensional Taylor series (with $o(S)$ denoting higher-order terms), and f_c, f_x in (1.6) to (1.18), F_u, G_u in (1.11), F_v, G_v in (1.18) are, of course, gradients. Keeping $n = m = 1$ simplifies the presentation and discussion of ideas. Generalizations for multidimensional x and/or u are beyond the scope of the paper; for $n = m = 2$, see [1, Ch. IX, Section 5]; for $n = 2$, $m = 1$, see examples in Section 8 below.

A similar to (1.5) functional equation, with slightly different derivation, is obtained in [1, Ch. IX, Section 13] for a nonautonomous problem where functions F, G of (1.3) depend explicitly on time $t \in [0, T]$. For this case, with fixed T, a nonstationary cost functional is introduced instead of (1.4), namely,

$$\max_u J^*(u) = \max_u \int_a^T F(x, u, t)\, dt = f(a, c), \tag{1.12}$$

where $c = x(a)$ is an intermediate point for variable $a \in [0, T]$. Similarly to (1.5), we have

$$f(a, c) = \max_{u[a,a+S]} \left[\int_a^{a+S} F(x, u, t)\, dt + f(a + S, c(a + S)) \right], \tag{1.13}$$

for all a and all S such that $0 \le a < a + S \le T$.

Now, the same operations as in (1.6)–(1.8) yield

$$f(a, c) = \max_{u[a,a+S]} \left[F(x, u, a)S + f(a, c) + Sf_a + Sf_c \frac{dc(S)}{dS} + o(S) \right], \tag{1.14}$$

$$0 = \max_{u(a)} [F(c, u(a), a) + f_a + f_c G(c, u(a), a)], \tag{1.15}$$

$$-f_a = \max_v [F(c, v, a) + f_c G(c, v, a)], \tag{1.16}$$

and for optimal $v = v^o$ in the interior of the region, we get, cf. [1, Ch. IX, Section 13, equation (8)],

$$-f_a = F(c, v^o, a) + f_c G(c, v^o, a), \tag{1.17}$$

$$0 = F_v(c, v^o, a) + f_c G_v(c, v^o, a). \tag{1.18}$$

Solving this system for f_a, f_c and equating $f_{ac} = f_{ca}$ yields a partial differential equation for $v^o(a, c)$ which is the extremal control $u^o(t, x(t))$.

REMARK 1.3. Formulae (1.13)–(1.18) for a nonautonomous problem can be derived in the same way as (1.5)–(1.11) for an autonomous problem by including time coordinate t into expanded space R^{n+1} as follows. Keeping $n = 1$ and setting $a = 0$ to fit the framework of (1.3)–(1.8), denote $(x, t) = y \in R^2$, so that, instead of (1.12), we consider equivalent autonomous representation, cf. (1.3),(1.4),

$$f(c, T) = \max \int_0^T F(y, u)\, d\tau, \quad \frac{dy}{d\tau} = [G(y, u), 1], \quad y(0) = [c, 0], \qquad 0 \le \tau = t \le T. \tag{1.19}$$

Expressions (1.5) remain unchanged with the substitution of (y, τ) for (x, t) but in (1.6) one should account for a new variable $t(\tau)$ and replace the middle term as follows:

$$Sf_c \frac{dc(S)}{dS} := S\, \nabla f \frac{dy(S)}{dS} = Sf_c G(c, 0, v) + Sf_t(0). \tag{1.20}$$

Now, $f(c, T)$ in (1.6) cancels out as in (1.7), and after division by S, we have, to the first order as $S \to 0$,

$$f_T = \max_v [F(c, 0, v) + f_c G(c, 0, v)] + f_t(0) \tag{1.21}$$

in place of (1.8). If T is fixed as in [1, Ch. IX, Section 13], then $f_T = 0$ and (1.21) is identical to (1.15),(1.16) for the starting moment $a = 0$, since $f_a = f_t(0)$ at the starting point $x[a = 0] = c$

does not depend on $u(\cdot)$. By the principle of optimality, for an optimal policy $u[0,T]$, the remaining decisions $u[a,T]$, $a > 0$ must constitute an optimal policy with regard to the resulting state $y(\tau)$, $\tau = a > 0$, and hence, the above derivation holds for any $a \geq 0$.

REMARK 1.4. If the extremal (supposed to be optimal) u^o of (1.9), or v^o of (1.16), is not in the interior of the admissible region, then PDEs (1.10),(1.11), or (1.17),(1.18), must be replaced by extremality (stationarity) conditions of the Karush-Kuhn-Tucker type. Moreover, in this case the classical theorem on mixed partial derivatives to find u^o, v^o may be inapplicable. In engineering and economics, optimal policies (controls) usually belong to boundaries of admissible regions for controls and states (trajectories), which necessitates the replacement of those PDEs (not the Bellman equations (1.9),(1.16)). This extension is beyond the scope of the paper.

There is vast literature on the subject, see, e.g., [1,3–9] and references therein, mainly concerned with the application of the optimality principle to various kinds of problems, derivation of the corresponding functional equations, their solution by different methods under certain assumptions and conditions, investigation of the existence of solutions and stability properties, numerical studies and computational experiments, comparison with other optimization methods, etc.

However, the principle of optimality is an important statement in its own right and it is interesting to study it as such, without an accent on its numerous applications, particular functional equations, technical details, or connections with other theories.

2. GENERALITY OF THE STATEMENT

The statement of the principle of optimality does not specify a notion of optimality to which the principle is referred. In [1], it is applied to max/min optimality for certain classes of functionals such as in (1.1)–(1.3) and (1.12), and to optimality in games with a saddle point. There are, however, other notions of optimality: multiobjective (Pareto) optimality, Nash optimality, optimality in games without a saddle point, local or global optimality, and vast variety of classes of functionals within each kind of optimality.

The statement does not specify a set of possible decisions (admissible controls), nor a class of problems to which it is applicable, nor allowable restrictions, etc., leaving one to believe that it is a universally valid principle, i.e., "a comprehensive and fundamental law, doctrine, or assumption" (Webster's dictionary), under the sole hypothesis that an optimal policy exists. It also does not require convexity, monotonicity, differentiability, etc., though such properties may be imposed on a case-to-case basis in order to obtain or solve functional equations resulting from the principle and to assure that an extremal (stationary) solution is, in fact, optimal.

In current literature, the principle is used mainly as a heuristic applied to various problems as long as it is possible to get the results. Such a pragmatic approach is reasonable, and to support it, a study to establish the overall validity and limitations of the principle of optimality is of paramount importance.

3. OBSERVATIONS

The formulation of the principle as presented in [1, p. 83], see Section 1 above, contains some explicit and implicit assumptions.

ASSUMPTION 3.1. *The first explicit assumption is the existence of an optimal policy which includes optimal remaining decisions as its parts.*

ASSUMPTION 3.2. *The second explicit assumption is the existence of "the state resulting from the first decision". Though the notions of "initial decision" and "the first decision" are not specified, it is clear from the context that they are identical and denote the whole part of "an optimal policy" from "the initial state" until the state from which "the remaining decisions"* *begin. This state is thus an arbitrary section (compare Dedekind sections in the set of real numbers) that separates "initial (first) decision" from "the remaining decisions". The set of such states (sections) represents a trajectory (not mentioned in the formulation of the principle) which may be called optimal as corresponding to an optimal policy. This trajectory with the initial state adjoined is called "the process" in the sentence preceding the formulation of the principle. It may contain discrete and continuous parts but it must not contain infinite collections of discrete states or discontinuities over any finite time period in the process. An optimal trajectory may consist of a finite number of different branches, all emanating from a single initial state. Each branch may be piecewise continuous (i.e., having a finite number of discrete states and/or discontinuities over any finite time period of the process) and generated by piecewise continuous decisions (controls).*

ASSUMPTION 3.3. *The first implicit assumption is the existence of nonoptimal policies and corresponding trajectories emanating from the same initial or intermediate state along an optimal trajectory. Indeed, if there are only optimal policies (e.g., for a constant cost function), the principle becomes a tautology.*

ASSUMPTION 3.4. THE CONGRUENCE PROPERTY. *For simplicity, suppose first that an optimal policy is unique and defined by a sequence of decisions* $\{u_1, \dots, u_n, \dots, u_N\}$ *or by a continuous function* $u(t)$, $0 \le t < T$, $T \le \infty$. *Then, "the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision". Denote this subsequent optimal policy by the sequence* $\{v_n, \dots, v_N\}$, $1 < n < N$, *or, respectively, by the function* $v(t)$, $0 < t_1 \le t < T$, *with the understanding that* n *or* t_1 *corresponds to the state resulting from decisions* $\{u_1, \dots, u_{n-1}\}$, *or* $u(t)$, $0 \le t < t_1$. *The congruence property is the requirement of the principle of optimality that* $v_i = u_i$, $i = n, \dots, N$, *or, respectively,* $v(t) \equiv u(t)$, $t_1 \le t < T$. *Indeed, otherwise,* $\{u_1, \dots, u_N\}$, *or* $u(t)$, $0 \le t < T$, *would not constitute an optimal policy as specified in the principle, if an optimal policy is unique. This can be taken as an immediate proof by contradiction which, indeed, appeared in the literature, see Section 4.*

If an optimal policy is not unique, then u_i $(i = 1, \dots, N)$ *and* $u(t)$ *above may be set-valued. In this case, the congruence property means that* $v_i \subseteq u_i$, $i = n, \dots, N$, *or* $v(t) \subseteq u(t)$, $t_1 \le t < T$. *For this and other reasons, see Section 6, we prefer to consider congruence as a separate assumption meaning the inclusion of "remaining decisions" in the remaining part of the original policy, without requirement of their optimality. There are different terms expressing this notion in the literature, e.g., "joining controls" in [7, p. 83], or "concatenated" pieces of admissible trajectory in [8, p. 87].*

ASSUMPTION 3.5. ADDITIVITY OF COSTS IN THE OPTIMALITY CRITERION (COST FUNCTION). *This assumption can be seen from the functional equations (1.1), (1.5), (1.13), and other similar equations derived from the principle of optimality. In some sources, e.g., in [7, pp. 84–85], it is postulated separately in order to prove the principle, see Section 4. In integral criteria, it follows from the additivity of the integral with respect to intervals of integration. This assumption represents a restriction on the class of problems to which the principle of optimality is applied.*

ASSUMPTION 3.6. FEEDBACK STRUCTURE OF OPTIMAL POLICIES. *This property is implied by the statement that "the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision". It can be seen from equations (1.1),(1.2) that define a policy* $q(p)$, *from (1.8) for* $v(c)$, *from (1.10),(1.11) for* $u^o(x)$, *from (1.17),(1.18) for* $v^o(a, c) \Rightarrow u^o(t, x(t))$, *and from other similar equations in the literature.*

ASSUMPTION 3.7. INDEPENDENCE ON THE PRECEDING DATA AND/OR DECISIONS. *The optimality of remaining decisions "with regard to the state resulting from the first decision" must not explicitly depend on the initial state nor the initial decision. This requirement is manifested by the word "whatever" in the formulation of the principle. In practice, it means that feedback controls obtained from equations implied by the principle explicitly depend on the current state, and not on the initial state nor preceding states, nor preceding decisions.*

ASSUMPTION 3.8. INVARIANCE OF THE BELLMAN FUNCTION $f(\cdot)$. *This hypothesis is not mentioned in the principle itself as formulated in [1, p. 83], see Section 1, and it does not follow from the principle nor from any of the preceding Observations 3.1–3.7. However, it is used in the derivation of the functional equations and PDEs of dynamic programming, see (1.2), (1.5), (1.6), (1.13), (1.14) at left and in the bracket, and it is tacitly assumed in applications and research works that followed in subsequent publications. This hypothesis is supported by a number of solved examples, though without any theoretical justification. The invariance of the function $f(\cdot)$ may be violated with the appearance of additional arguments which are not present in the original Bellman's procedures, see Example 6.1, Case 1.*

4. PROOF BY CONTRADICTION

Numerous problems solved with the help of dynamic programming dispel any doubt of its practical utility. This, however, does not answer reasonable questions about universality and general validity of the principle of optimality as a necessary condition.

Geometrically, the principle of optimality can be illustrated as follows. Consider a period of time $[t_o, T]$, $T \leq \infty$, an admissible region X (state space), and denote by $x(t) \in X$ a point on a trajectory which may depend on a policy (control) $u(x,t)$ applied to the process to assure the optimality of some criterion $J(u(\cdot), \cdot)$, say, $\max J$, see [1,3]. A trajectory $x[t_o, T)$ can be discrete or continuous. Admissible controls are piecewise continuous functions, and the cost function (optimality criterion) $J(u(\cdot), \cdot)$ is assumed to be additive [1,3–9], that is

$$J[t_1, t_3] = J[t_1, t_2] + J[t_2, t_3], \qquad t_0 \leq t_1 < t_2 < t_3 \leq T. \tag{4.1}$$

The policy (control)

$$u[t_1, t_2] = u^\circ(x,t) = \arg\max J(u, \cdot), \qquad t \in [t_1, t_2], \tag{4.2}$$

and the corresponding piece of trajectory $x[t_1, t_2]$ is called optimal. The statement of the principle is simple: *if $u[t_o, T)$ is optimal for the whole trajectory $x[t_o, T)$, then remaining parts $u[t, T)$, $x[t, T)$ are also optimal for any $t \in (t_o, T)$, provided that $x(t) \in x[t_o, T)$. Shortly: remaining parts of an optimal policy and trajectory are themselves optimal with respect to the same criterion.*

In some sources, see, e.g., [9, footnote to (4.2.16)], this condition is qualified as a sufficient condition. Other sources present a proof of the principle as a necessary condition of optimality. In [7, pp. 86–87], or [8, p. 87], a simple proof of the principle is given which can be summarized as follows.

PROOF. Setting $t_1 = t_o$, $t_3 = T$ in (4.1) and assuming the existence of optimal policies (trajectories) starting at any point $x(t) \in x[t_o, T)$, one can see that if $J[t_o, T] = J^\circ$ is optimal, say, $J^\circ = \max J[t_o, T]$ for some $u^\circ(t)$, $t \in [t_o, T)$, then it implies

$$\max J[t_o, T] = J[t_o, t_2] + \max J[t_2, T]. \tag{4.3}$$

Indeed, if $J[t_2, T]$ in (4.3) is not maximal, then there is a better policy $u[t_2, T)$, and thus, $J[t_o, T]$ at the left in (4.3) is not optimal, contradicting the assumption of its optimality.

Note that this proof does not use Invariance Hypothesis 3.8, nor important Observations 3.6, 3.7 of Section 3.

5. TOTAL OPTIMALITY

With the same idea, a much stronger result can be proved.

THEOREM 5.1. *If for an optimal trajectory, all remaining decisions are optimal, then every part of that trajectory is itself optimal.*

PROOF. Consider relation (4.1) again with $t_3 = T$, fixing neither t_1 nor t_2. By assumption, $J[t_1, T]$, $J[t_2, T]$ are optimal, say, maximal, as corresponding to optimal remaining decisions, and thus, instead of (4.1), we can write

$$\max J[t_1, T] = J[t_1, t_2] + \max J[t_2, T], \qquad t_o \leq t_1 < t_2 \leq T. \tag{5.1}$$

Now, if $J[t_1, t_2]$ in (5.1) is not maximal, there is a better policy $u[t_1, t_2]$, and thus, $J[t_1, T]$ at the left in (5.1) is not optimal, contradicting the assumption of its optimality.

REMARK 5.1. In the above two proofs, we encounter the plausibility phenomenon that lies at the root of intuition. Indeed, the statement "if a trajectory is optimal, then all its parts are also optimal" is unbelievable. By the way, it is correct for free autonomous systems with respect to certain notions of optimality, see Section 6 below. In contrast, the statement "if a trajectory is optimal, then all its remaining parts are optimal" is highly believable. If we replace the word "trajectory" with the word "policy", then both statements seem questionable. That is why the note "A proof by contradiction is immediate" in the citation was needed (Section 1). In fact, relations (4.3) and (5.1) demonstrate that the two statements are equivalent. If for some reason the proof with (5.1) is discarded, then the proof with (4.3) must also be discarded, for the same reason, so that the proofs of Bellman's principle of optimality presented in the literature would disappear, leaving one to accept the principle by virtue of a large number of solved examples, that is, by popular belief, or "by definition".

REMARK 5.2. The existence of totally optimal trajectories is readily seen without any optimality arguments. Indeed, if a problem

$$f(c, a, T) = \min_u \int_a^T [F(x, c, u, t)]^2 \, dt, \quad \frac{dx}{dt} = G(x, u, t), \qquad t \geq 0, \quad x(0) = c, \tag{5.2}$$

is controllable, i.e., $\frac{\partial G}{\partial u} \neq 0$ almost everywhere, then the global $\min f(\cdot) = 0$ if and only if $F(x, c, u, t) = 0$ on trajectories of the system for all t. If there exists a control $u^o(x, t)$ respecting this constraint, i.e., $F(x, c, u^o(x, t), t) \equiv 0$ on trajectories of the equation $\frac{dx}{dt} = G(x, u^o(x, t), t)$, $x(0) = c$, then those trajectories are totally optimal in the sense that along every part of every trajectory we have $f(c, a, T) = 0$ for any c, a, T. Examples: $F_1 = x - ce^{-t}$, $F_2 = x + u$, $\frac{dx}{dt} = u$, $x(0) = c$; then with $u^o = -x$ we have $f(c, a, T) = 0$ on any interval $[t^*, T)$, $T > a \geq t^* \geq 0$, for both F_1 and F_2. For another illustration, with $F(x, u, t)$, see Example 6.1, Case 4, below.

REMARK 5.3. It is problematic to use the optimality of intermediate parts to try to simplify the structure of functional equations in dynamic programming. Indeed, even if the first term in (1.5) is optimal by itself, one cannot open the brackets in (1.5) since $c(S)$ depends on the choice of $u[0, S)$, and the same for (1.13).

COROLLARY 5.1. *If two remaining parts $x[t_1, T]$, $x[t_2, T]$ of a trajectory (not necessarily optimal) are optimal, then the piece $x[t_1, t_2]$ of that trajectory is also optimal.*

DEFINITION 5.1. *If all parts of a trajectory including the whole trajectory itself are optimal, then such a trajectory (process) is called totally optimal.*

DEFINITION 5.2. *If all parts of a trajectory including the whole trajectory itself have the same extremality (stationarity) conditions, then such a trajectory (process) is called totally extremal.*

With this definition, the principle of optimality can be applied to stationary trajectories (extremals), see Remark 1.1 above and Example 6.1 below. The proofs in (4.3),(5.1) remain the same, and in Theorem 5.1 the term "optimal" should be replaced by the term "extremal".

We see that the principle of optimality under Assumptions 3.4–3.8 implies total optimality. It is well known that certain optimal autonomous systems are totally optimal. Certain optimal nonautonomous systems may also be totally optimal. Clearly, every totally optimal system satisfies the principle of optimality.

6. TOTAL OPTIMALITY IN DYNAMICAL SYSTEMS

Consider an autonomous system in (1.3) with fixed T

$$\max/\min J(u) = \int_0^T F(x,u)\,dt, \qquad \frac{dx}{dt} = G(x,u), \qquad \frac{\partial G}{\partial u} \neq 0, \qquad x(0) = c. \tag{6.1}$$

The total extremality of (6.1) follows, of course, from (1.10),(1.11) with $f_T = 0$, see [1, Ch. IX, Section 3, equation (14)]. However, for solvability of determinantal equation in terms of u^o in (1.10),(1.11), the condition $\frac{\partial G}{\partial u} \neq 0$ is insufficient. Also, relations (1.10),(1.11) are derived from the optimality principle which is equivalent to total optimality. To avoid this vicious circle, let us prove the total optimality of (6.1) with the help of variational calculus under the assumption that the functional $J(u)$ is strictly locally concave for max, or convex for min, in a neighborhood of an extremal trajectory.

PROOF. Using the inclusion lemma [10, pp. 723–724], problem (6.1) of optimal control can be converted into the problem of the calculus of variations

$$\max/\min \int_0^T F\left(x, h\left(x, \frac{dx}{dt}\right)\right) dt, \qquad x(0) = c, \tag{6.2}$$

where u is eliminated, due to the existence of certain unique function $u = h(x, \frac{dx}{dt})$ for which $\frac{dx}{dt} \equiv G(x, h(x, \frac{dx}{dt}))$, see [10]. With the substitution $u = h(x, \frac{dx}{dt})$, the integrals in (6.1),(6.2) have the same values on trajectories of the equation $\frac{dx}{dt} = G(x,u)$ for any initial conditions, and hence, problems (6.1),(6.2) are equivalent.

The Euler-Lagrange equation for a stationary curve of (6.2) is of the form

$$F_x - \frac{dF_{x'}}{dt} = F_x - x'F_{xx'} - x''F_{x'x'} = 0, \qquad x' = \frac{dx}{dt}, \tag{6.3}$$

and if a solution exists, then (6.3) has the first integral

$$F(x,x') - x'F_{x'}(x,x') = C = \text{const} \tag{6.4}$$

that yields a one-parameter family of extremals, each of which is defined by two conditions: either initial conditions $x(0) = c$, $x'(0) = G(c,v)$ where $v = u^o(0)$ for yet unknown optimal policy, (control) u^o, cf. (1.7),(1.8) in Bellman's derivation, or boundary conditions $x(0) = c$, $x(t_1) = c_1$ (sic: controllability condition should be satisfied for $x(t_1) = c_1$). Since C in (6.4) is an arbitrary constant that can be chosen to fit any condition $x(t_1) = c_1$, it is clear that (6.3),(6.4) are fulfilled on any part of a stationary curve between $x(t_0) = c_0$, $x(t_1) = c_1$, $0 \leq t_0 < t_1 \leq T$, whence it follows that any part of an extremal is itself an extremal. Now, if an extremal is, in fact, an optimal trajectory on $[0,T]$, then it follows that every part of it is also optimal. Hence, an optimal autonomous system (6.1) is totally optimal. Note that the function $F(\cdot)$ in (6.2) is supposed to be continuously differentiable up to the third order with respect to all its arguments. There is no need to check the Legendre or Jacobi necessary conditions: if they are fulfilled for the whole optimal trajectory, they are fulfilled for any part thereof.

The variables x, u in (6.1) can be considered as vectors of different dimensions: $x \in R^n$, $u \in R^m$. Then in (6.1), and in the inclusion lemma [10], the condition $\frac{\partial G}{\partial u} \neq 0$ should be replaced by the full rank condition of the Jacobian: $\operatorname{rank}(\frac{\partial G}{\partial u}) = \min(n,m)$, with possible switching between its rows. Then there will exist certain functions $u_j = h_j(x,x')$, $j = 1,\ldots,m$, with possible switching for different $h_j(\cdot)$ on different parts of trajectories (if $m > n$, there are redundant controls that may be used to assure the full rank condition for $\frac{\partial G}{\partial u}$).

With those $h_j(\cdot)$, problem (6.1) is equivalent to a collection of problems (6.2), or to just one problem (6.2). For each problem (6.2), we have the same n equations (6.3) and n first integrals (6.4) for x_i, C_i, $i = 1, \dots, n$. By the same argument as above, it follows that optimal autonomous problems (6.1) with $x \in R^n$, $u \in R^m$, $\operatorname{rank}(\frac{\partial G}{\partial u}) = \min(n, m)$ are totally optimal.

REMARK 6.1. Clearly, a boundary condition or some autonomous constraints can be added to (6.1), transforming (6.2) into a Lagrange problem that may have an optimal solution. Though still autonomous and optimal (in another sense), this new problem may lose the quality of total optimality of its trajectories and controls, see Example 8.2 below.

For nonautonomous systems (6.1) where either F, or G, or both depend explicitly on time t, the Euler-Lagrange equation is of the form

$$\frac{d\,[F - x'F_{x'}]}{dt} - F_t = 0, \tag{6.5}$$

and if a solution exists, then (6.5) has the first integral on $[t_0, T] \subseteq [0, T]$,

$$F(t, x, x') - x'F_{x'}(t, x, x') = C + \int_{t^0}^{t} F_t(t, x, x')\, dt, \tag{6.6}$$

where for simplicity we used the same notation $F(t, x, x') := F(t, x, h(t, x, \frac{dx}{dt}))$. If $F_t \equiv 0$, then (6.6) coincides with (6.4). Comparing (6.6) with (6.4), we see that, in general, nonautonomous optimal systems are not totally optimal except for some special cases, cf. Remark 5.2.

REMARK 6.2. By formal consideration of time t as a new state variable with the addition of the special auxiliary equation $\frac{dt}{d\tau} = 1$, $t(0) = 0$, see Remark 1.3, any nonautonomous system $x(t) : R \to R^n$ can be formally represented as an autonomous one in the time-space R^{n+2}, one dimension greater than initial time-space R^{n+1}. However, this new "autonomous" system evolves not in R^{n+2} but in its plane section $t = \tau$, that is, still in R^{n+1}. In light of Remark 6.1, one cannot expect to gain total optimality by such considerations.

EXAMPLE 6.1. Consider the class of problems with $x(t) \in R$ and appropriate $F(z)$,

$$\begin{gathered} f = \min/\max \int_0^T F(r_1 t + r_2 x + r_3 u)\, dt, \qquad x' = \frac{dx}{dt} = x^n u^r + bt, \\ n \geq 0, \qquad x(0) = c. \end{gathered} \tag{6.7}$$

CASE 1. If we take max in (6.7) with linear $F(z) = z$, and also $r_1 = n = b = 0$, $r_2 = -r_3 = 1$, $r = 0.5$, the problem is autonomous and coincides with the example in [1, Ch. IX, Section 10] where a supplementary constraint is considered

$$0 \leq u(t) \leq x(t), \qquad t \in [0, T], \tag{6.8}$$

that reflects a process of optimal resource allocation, see constraint (6.8) also in [1, Ch. IX, Sections 3,7,10–12] and [3, Ch. 3, Sections 11,12; Ch. 6, Sections 1,3,4; Ch. 8, Section 2; Ch. 9, Sections 1–3,6,7,12,14].

VARIATIONAL SOLUTION. Eliminating $u = x'^2$ from (6.7) and considering variable gain period T, we obtain the collection of equivalent variational problems

$$f(c, T) = \max \int_0^T \left(x - x'^2\right) dt, \qquad x(0) = c, \tag{6.9}$$

which lack the endpoint condition $x(T)$ and are subject to constraint (6.8): $0 \leq x'^2 \leq x$ for all $t \in [0, T]$, which we omit for the moment.

The Euler-Lagrange equation (6.3) for problem (6.9) has the form $1+2x''=0$ with the solution

$$x(t) = -0.25t^2 + c_o t + c, \qquad c_o = x'(0) \geq 0, \tag{6.10}$$

where c_o is not specified but subject to the constraint $u(0) = c_o^2 \leq c = x(0)$, and thus, to have a nontrivial problem in (6.9), we must consider $c > 0$. Since $F_{x'x'} = -2 < 0$, so every curve (6.10) is optimal and renders max in (6.9) with respect to all neighboring curves. Simple calculation yields optimal control

$$u^o = x'^2 = c_o^2 + c - x, \qquad c > 0 \text{ given}, \tag{6.11}$$

and since c_o is not specified, we have a one-parameter family of optimal trajectories, each corresponding to a particular linear feedback of (6.11). Clearly, every trajectory is totally optimal, so that we have a totally optimal field of trajectories depending on two parameters c, c_o, $c \geq c_o^2$.

Over this field of optimal trajectories, integral (6.9) takes the values

$$f(c,T,c_o) = \int_0^T \left(2x - c_o^2 - c\right) dt = \left(c - c_o^2\right) T + c_o T^2 - \frac{T^3}{6}, \tag{6.12}$$

which form a surface in R^4. For a fixed $T > 0$, the maximum maximorum of the performance for appropriate choice of c_o is

$$\max_{c_o} f(c,T,c_o) = cT + \frac{T^3}{12}, \qquad \text{at } c_o^* = 0.5T. \tag{6.13}$$

From (6.8), (6.10), (6.11), we have the inequalities

$$0 \leq u^o(t) = c_o^2 + c - x(t) \leq x(t) = -0.25t^2 + c_o t + c, \qquad t \in [0,T], \tag{6.14}$$

which imply limits on the lifetime T and on the optimal rate of investment $u^o(t)$,

$$0 \leq t \leq T \leq \max T = T^* = 2c_o + 2\left(c_o^2 + c\right)^{1/2}, \tag{6.15}$$

$$0 \leq u^o(t) = c_o^2 + c - x(t) \leq x(t) \leq \max x(t) = c_o^2 + c, \qquad \text{at } t = 2c_o < 0.5T^*. \tag{6.16}$$

Let us consider the gain/loss intervals in (6.9) for fixed $c_o < c^{1/2}$ which depend on the sign of the integrand in (6.9). We have for the gain intervals in (6.9),(6.12),

$$x - x'^2 = 2x - c_o^2 - c = -0.5t^2 + 2c_o t + c - c_o^2 \geq 0. \tag{6.17}$$

The roots are

$$t_{1,2} = 2c_o \pm \left(2c_o^2 + 2c\right)^{1/2}, \qquad t_1 \leq 0 < t_2, \quad \text{since } c_o^2 \leq c, \tag{6.18}$$

and we have

$$x - x'^2 \geq 0, \qquad \text{if } t \in [0, t_2]; \qquad x - x'^2 < 0, \qquad \text{if } t \in (t_2, T^*], \tag{6.19}$$

and thus, $[0, t_2)$ is the gain period, $(t_2, T]$ is the loss period, so the cut-off time is

$$t_2 = 2c_o + \left(2c_o^2 + 2c\right)^{1/2} < T^* = 2c_o + 2\left(c_o^2 + c\right)^{1/2}. \tag{6.20}$$

For example, if $c = 300$, $c_o = 10$, then we have $T^* = 60$, $t_2 \cong 48$. If we take $T = 48$, then with the same $c_o = 10$, we have, by (6.12), $f(300, 48, 10) = 14208$. Since $c_o \leq c^{1/2} = 17.3$, so for the optimal pair in (6.13) with $T = 34$, $c_o^* = 17$, we have $f(300, 34, 17) = 13475$ by (6.12) and by (6.13), which is 5% less than $f(300, 48, 10)$.

DYNAMIC PROGRAMMING APPROACH. For the problem

$$f(c,T) = \max \int_0^T (x-u)\,dt, \qquad \frac{dx}{dt} = u^{1/2}, \qquad x(0) = c, \tag{6.21}$$

the PDEs (1.10),(1.11) are at $t = 0$, denoting $v = u(0)$,

$$f_T = c - v + f_c v^{1/2}, \qquad 0 = -2 + f_c v^{-1/2}, \tag{6.22}$$

whence

$$f_c = 2v^{1/2}, \qquad f_T = c - v + 2v = c + v, \qquad f(c,0) = 0. \tag{6.23}$$

For $t > 0$, we have to substitute $x(t)$ for c, and $u(t)$ for v, which, together with condition (6.8), yields

$$f_x = 2u^{1/2}, \qquad f_T = x + u, \qquad 0 \le u \le x. \tag{6.24}$$

Finding extremal feedback $u^o(x,t)$ by the theorem on mixed partial derivatives is impossible since $f_{xT} = 0$, $f_{Tx} = 1$. If we fix T, then $f_T = x + u = 0$, so that $u = -x$, and for $x(t) > 0$, the expression $u^{1/2} = (-x)^{1/2}$ has no real values. If, to bypass this difficulty, we consider in (6.21) the equation $(\frac{dx}{dt})^2 = u \ge 0$, then $G(x,u) = u^{1/2}$ disappears, and equations (1.10), (1.11), (6.22)–(6.24) do not make sense. These difficulties preclude the application of PDEs of dynamic programming for solution of problem (6.21), and they necessitated in [1, Ch. IX, Section 10] the return to the calculus of variations to solve this problem.

Comparing the dynamic programming approach with variational solution, one can see that invariance of the Bellman function $f(c,T)$ in (6.21) is violated. Indeed, the postulated cost function in (6.9) acquired additional argument c_o in (6.12),(6.13) that affected not only the optimal control (6.11) but also the gain/loss periods (6.19),(6.20) which are completely lost in dynamic programming procedure due to the imposition of rigid structure in (6.21)–(6.24) that led to the contradiction $f_{xT} \ne f_{Tx}$ (note that in (6.12) we have $f_{cT} = f_{Tc} = 1$). What is worse, the absence of an argument c_o is not clear from dynamic programming procedure; in contrast, this argument c_o appears naturally from the Euler-Lagrange equation, see (6.10)–(6.12).

Such difficulties of dynamic programming may appear despite total optimality of all optimal trajectories (in which case the principle of optimality holds), and despite the existence of optimal controls in the form of a feedback, as in (6.11). They are exposed here not to diminish the importance of dynamic programming, but to bring attention of unfamiliar users to possible intricacies of the method.

CASE 2. $n = 0$, $r_1 = r_2 = r_3 = r = b = 1$, $\frac{\partial G}{\partial u} = 1$, so we have the equivalent variational problems

$$f = \min/\max \int_0^T F(x + x')\,dt, \qquad x(0) = c, \qquad x' = \frac{dx}{dt}. \tag{6.25}$$

Since $F(x + x')$ does not depend explicitly on t, we see that, in this case, nonautonomous control problems (6.7) have autonomous variational representation (6.25) without any constraints. In such cases, nonautonomous optimal systems are totally optimal, and thus, principle of optimality holds on optimal trajectories of such systems.

Let us take $F(z) = z$ in (6.25). Then, there are no optimal (min or max) solutions. However, stationary (extremal) solutions do exist.

VARIATIONAL SOLUTION. Equation (6.4) yields

$$x + x' - x' = x = C = \text{const} = c, \tag{6.26}$$

for which (6.3),(6.5) become trivial identities, and from (6.25) we have $f = cT$, and $u^o = -t$ from (6.7) for this case. The extremals (Lagrange curves) are given by the family of horizontal lines, and extremal control is not a feedback. With $F(z) = z$, the functional in (6.7) is linear, and thus, convex and concave but not strictly, and we see that the extremal policy $u^o = -t$ is totally extremal but not optimal. If we consider in (6.7) the min with $F(z) = z^2$, this policy is still totally extremal but not optimal. However, in this case, there is a family of solutions: $x = c_1e^t + c_2e^{-t}$, $c_1 + c_2 = c$, which are totally optimal (locally, $F_{x'x'} = 2 > 0$), with $f = 2c_1^2(e^{2T} - 1)$. If $c_1 = 0$, $c_2 = c$, then $f = 0$ for $x = ce^{-t}$ with $u^o = -x - t$ (global optimal solution).

DYNAMIC PROGRAMMING APPROACH. With $F(z) = z$, equations (1.17),(1.18) yield for the integral over $[a, T] \subseteq [0, T]$ in (6.7) with $r_1 = r_2 = r_3 = r = b = 1$, $n = 0$,

$$-f_a = a + c + v + f_c(v + a), \tag{6.27}$$

$$0 = 1 + f_c, \tag{6.28}$$

and thus, $f_c = -1$, $f_a = -c$, and v is not defined. By (6.7), $f_a = -F[a + x(a) + u(a)] = -c$, so

$$f = \int_0^T F(\cdot)\, dt = \int_0^T c\, dt = cT, \qquad \text{same.} \tag{6.29}$$

In this case, dynamic programming equations (6.27),(6.28) fail to directly define an extremal policy despite its existence and the validity of the principle of optimality specified for the extremal policy which can be determined as follows: by (6.7),(6.29), we have $t + x + u = c = \text{const}$, and thus, $\frac{du}{dt} = -x' - 1 = -(u + t) - 1$, with $u = -t$ as an extremal control. For min with $F(z) = z^2$, equation (1.18), if we set $f_c = 0$, yields the equation $a + c + v = 0$, or $x + x' = 0$, with global optimal solution $x = ce^{-t}$. Otherwise, the equality $f_{ca} = f_{ac} \neq 0$ yields the extremal $x(t) \equiv 1$.

CASE 3. $r_1 = r_2 = r_3 = r = 1$, $b = 0$, $\frac{\partial G}{\partial u} = x^n \neq 0$, $u = x'x^{-n}$. For $F(z) = z$, the equivalent variational representation

$$f = \min/\max \int_0^T \left(t + x + x'x^{-n}\right) dt, \qquad x(0) = c, \qquad x(t) \neq 0, \tag{6.30}$$

is nonautonomous. However, if we drop t from (6.30), it becomes autonomous with the addition of $0.5T^2$ to the cost function. We see that in some cases a nonautonomous extremal problem can be reduced to an autonomous one such that its extremal trajectories are totally extremal. This can be seen from (6.6) which for (6.30) takes the form

$$t + x + x'x^{-n} - x'x^{-n} = C + t - t_0, \qquad x = C - t_0 = \text{const} = c, \tag{6.31}$$

where terms with t, x' cancel out. With $t_0 = 0$, we get an extremal solution $x(t) = c$, $u^o = 0$, $f = cT + 0.5T^2$.

CASE 4. Global optimal solution. Consider min in (6.7) with $r_1 = r_2 = r_3 = r = 1$ and $F(z) = (z^2)^p \geq 0$, $p > 0$. Then, it is clear that global minimum is delivered by the optimal control $u^o = -x - t$ with the process equation $\frac{dx}{dt} = -x^{n+1} + (b - x^n)t$, $x(0) = c$, for which $f = 0$. If $n = 0$, $b = 2$, this yields $\frac{dx}{dt} = -x + t$ with the globally optimal trajectory $x(t) = t - 1 + (c + 1)e^{-t}$ which is also totally optimal. It is interesting that this solution is obtained from (1.17),(1.18) for $F(z) = z^2$ if we set $f_c = 0$. Indeed, we have for $n = 0$, $b = 2$,

$$-f_a = (c + v + a)^2 + f_c(v + 2a), \qquad 0 = 2(c + v + a) + f_c, \tag{6.32}$$

yielding $x + u + t = 0$ if $f_c = 0$, $f_a = 0$. However, if $f_c \neq 0$, then, according to the general procedure, we equate $f_{ac} = f_{ca} = -2$, yielding $-2(c + v + a) + 2(v + 2a) = -2$, so that the solution $x = t + 1$, $u = -2t + 1$, $f = 4T$ is obtained which exists only for $x(0) = 1$, with no neighboring solutions at all.

Example 6.1 demonstrates that nonautonomous systems may sometimes be totally optimal or reduced to totally optimal systems for which the principle of optimality is valid. Nevertheless, the PDEs of dynamic programming may, in some cases, not define a feedback control. Moreover, dynamic programming, in general, may not deliver a global optimal solution. It is important to understand, since the local character of the calculus of variations is clear from its title, which is not the case for dynamic programming nor for the max/min principle.

7. CONTIGUITY AND DYNAMIC PROGRAMMING

If an optimal trajectory (policy) is not totally optimal, then the principle of optimality is not valid for such a trajectory (policy). However, a modification of the derivation of the functional equations for the optimal cost and optimal policy allows one to conclude that the same resulting equations (1.16)–(1.18) hold under certain supplementary conditions, irrespective of the principle of optimality.

As in [1], we make a blanket assumption that all functions have continuous partial derivatives up to the order we may need, and that optimal trajectories we may consider do exist. In the course of derivation, we shall indicate all additional conditions that are required and must be verified when solving practical problems.

If the principle of optimality does not apply to the problem, then the transition from (1.12) to (1.13) analogous to (1.5a),(1.5b) is not valid. Moreover, an optimal cost $f(a,c)$ in (1.12) for the remaining policies on $[a,T)$, $a>0$, need not be preserved as the same function for all $t\in[a,T)$ because the optimal policy $u_t^o(\cdot)$ in the integrand may not be congruent with $u_a^o(\cdot)$, that is, $u_t^o(t)\neq u_a^o(t)$ for $t>a$, where $u_t^o(\cdot)$, $u_a^o(\cdot)$ are curves, not partial derivatives. With this in mind, denote

$$\max_u J^*(u) = \max_u \int_a^T F(x,u,t)\,dt = f^a(a,c), \tag{7.1}$$

where $c=x(a)$ is an intermediate point for variable $a\in[0,T)$. Now, we have

$$f^a(a,c) = \max_{u[a,T]}\left[\int_a^{a+S} F(x,u,t)\,dt + \int_{a+S}^T F(x,u,t)\,dt\right] \tag{7.2}$$

$$= \max_{u[a,a+S]}\left[\int_a^{a+S} F(x,u,t)\,dt + f^{a+S}(a+S,x(a+S)) + g^{a+S}(\cdot)\right], \tag{7.3}$$

where $g^{a+S}(\cdot)$ is the difference between the value of the second integral in (7.2) with the optimal control $u^o[a,T)$ and the optimal value $f^{a+S}(\cdot)$ of the same integral maximized separately by $u^*[a+S,T)$ which may not be a restriction of $u^o[a,T)$ onto $[a+S,T)$, that is, $u^*(t)\neq u^o(t)$ for $t\in[a+S,T)$ (incongruence), if the trajectory is not totally optimal.

THE WEAK INVARIANCE HYPOTHESIS. Assume that $\lim f^{a+S}(\cdot)=f^a(\cdot)$ as $S\to 0$ for every $a\in[0,T)$, and the same for its partial derivatives of the first order.

Now, expanding the first two terms of (7.3) into Taylor series with respect to S and denoting $v=u(a)$, we obtain

$$f^a(a,c) = \max_{u[a,a+S]}\left[F(c,v,a)S + f^a(a,c) + Sf_a^a(\cdot) + Sf_c^a(\cdot)\frac{dc(S)}{dS} + o(S) + g^{a+S}(\cdot)\right]. \tag{7.4}$$

In (7.4), we cancel out the term $f^a(a,c)$, then divide the equation by $S>0$, take $f_a^a(a,c)$ to the left-hand side, and consider $S\to 0$. Suppose that the following assumption holds:

$$\lim_{S\to 0} g^{a+S}(\cdot)/S = 0, \qquad \text{for all } a\in(0,T), \tag{7.5}$$

which we call the *contiguity condition*. Assume also that the Taylor series in (7.4) is convergent, and all terms in the bracket (7.3) with their respective partial derivatives are continuous in all arguments. Then, noting that $u(a)=v$, $x(a)=c$, we get the equation

$$-f_a^a(a,c) = \max_v\left[F(c,v,a) + f_c^a(a,c)G(c,v,a)\right], \qquad a\in[0.T), \tag{7.6}$$

and superscript a is redundant because of (7.5), thus, (7.6) is identical to (1.16). Since by notation $a = t$, $c = x(t)$, $v = u(a) = u(t)$, so relation (7.6) can be written in the usual form

$$-f_t(t,x) = \max_{u(t)}[F(x,u,t) + f_x(t,x)G(x,u,t)], \qquad t \in [0,T). \tag{7.7}$$

If optimal $u^o(t)$ is in the interior of the admissible region, then we have the equations

$$-f_t(t,x) = F\left(x,u^o,t\right) + f_x(t,x)G\left(x,u^o,t\right), \qquad t \in [0,T), \tag{7.8}$$

$$0 = F_u\left(x,u^o,t\right) + f_x(t,x)G_u\left(x,u^o,t\right), \qquad t \in [0,T), \tag{7.9}$$

which are identical to (1.17),(1.18). This means that the contiguity condition (7.5) is sufficient for derivation of those equations without the optimality of remaining trajectories. Vice versa, it is easy to see that, if the limit in (7.5) is nonzero or does not exist, equation (7.6) does not follow from (7.4), and thus, equations (7.8),(7.9) are invalid despite sufficient smoothness and convergence of the Taylor series in (7.4). Hence, under those regularity conditions, the following statement is proved.

THEOREM 7.1. *Under the Weak Invariance Hypothesis and the assumptions cited in Observations 3.5, 3.7, Section 3, the derivation of the functional equations and PDEs of dynamic programming is valid if and only if the contiguity condition (7.5) is satisfied, irrespective of the principle of optimality.*

REMARK 7.1. It is worth noting that for totally optimal systems (e.g., certain autonomous systems), we have $g^{a+S}(\cdot) \equiv 0$ for all a, S which means precisely the optimality of all remaining trajectories. Also, the validity of derivation may not imply the existence of solutions (same as indicated in [1] on many occasions).

Contiguity condition (7.5) can be simplified. To make it independent of the concrete problem above and special notations in (7.1)–(7.5), we return to considerations of Sections 4, 5, see (4.1), (4.3), (5.1).

In our notations of (5.1), the limit as $S \to 0$ in Section 13, Ch. IX of [1], corresponds to the limit as $t_2 \to t_1$, $t_1 \in [t_o, T)$ in (5.1). Without optimality of remaining decisions, we have to consider two different trajectories and, instead of (5.1), the following relations (the set-wise limit in (7.10) below means simple coincidence as $t_2 = t_1$, $t_1 \in [t_o, T)$, without consideration of various definitions of such limits):

$$x[t_1,T] = \lim_{t_2 \to t_1} \left(x[t_1,t_2] \cup x^*[t_2,T]\right), \quad x(t_2) = x^*(t_2), \qquad t_o \le t_1 < t_2 < T, \tag{7.10}$$

$$J(x[t_1,T]) = \lim_{t_2 \to t_1} \left\{J(x[t_1,t_2]) + \max J\left(x^*[t_2,T]\right)\right\}. \tag{7.11}$$

Here $x[t_1,T]$ designates a remaining (not necessarily optimal on $[t_1,T]$) part of an entire optimal trajectory $x^o[t_o,T]$, on which part the value $J(x[t_1,T])$ is attained under original optimal control $u^o[t_o,T)$. If we remove the equality signs at left and the limits in (7.10),(7.11), then $x[t_1,t_2]$ designates the initial part of semitrajectory $x[t_1,T]$ with the value $J(x[t_1,t_2])$ upon $x[t_1,t_2]$. Notation $x^*[t_2,T]$ means certain optimal semitrajectory starting at $x(t_2)$ with the value $\max J(x^*[t_2,T])$ attained on it. This semitrajectory $x^*[t_2,T]$ may not coincide with the remaining part $x[t_2,T]$ of the trajectory $x[t_1,T]$.

In cases when the principle of optimality is valid for the system under consideration, we have identically $x^*[t_2,T] \equiv x[t_2,T]$, equalities in (7.10),(7.11) are valid without the limit operation, relation (7.11) coincides with (5.1), and relation (7.10) becomes trivial tautology omitted in the proof of Theorem 5.1.

If the principle of optimality is invalid, then $x^*(t) \neq x(t)$, $t \in (t_2, T]$, but relations (7.10),(7.11) may still hold in the limit for any $t_1 \in [t_o, T)$. A careful reader can see that it is the equalities of type (7.10),(7.11) that are actually necessary and sufficient for the derivation of the functional equations and PDEs of dynamic programming in [1, Ch. 9, Sections 3,6,13,22], as it is presented above in (7.1) to (7.9).

Denote $J(x^\circ[t_2, T])$ the (nonoptimal) value of $J(\cdot)$ over remaining part $x^\circ[t_2, T]$ of the semitrajectory $x[t_1, T]$, $t_2 > t_1 = a \geq t_o$, under original optimal control $u^\circ(t) \in u^\circ[t_o, T]$. Let $\operatorname{extr} J = J(x^*[t_2, T])$ be the extremal (max or min) value of $J(\cdot)$ over a different semitrajectory $x^*[t_2, T]$ starting at the same point $x^*(t_2) = x^\circ(t_2)$ but optimized separately by some control $u^*(t) \neq u^\circ(t)$, $t \in [t_2, T)$.

Instead of $g^{a+S}(\cdot)$ in (7.3)–(7.5), consider the functional

$$C[t_1, t_2] := C\left(u^\circ[t_1, T], u^*[t_2, T]\right) = J\left(t_1, x^\circ[t_2, T]\right) - J\left(x^*[t_2, T]\right), \quad u^\circ[t_1, T] \subset u^\circ[t_o, T], \tag{7.12}$$

where t_1, t_2 are variable *markers* indicating two starting points of two incongruent remaining trajectories. The quantity $C[t_1, t_2]$ of (7.12) is identical to $g^{a+S}(\cdot) > 0$ for min-problem, and to $g^{a+S}(\cdot) < 0$ for max-problem as in (7.1)–(7.4). The value $J(t_1, x^\circ[t_2, T])$ is equal to the second integral in (7.2), and $J(x^*[t_2, T])$ equals the term $f^{a+S}(\cdot)$ in (7.3).

The functional (7.12) is a measure of J-proximity of semitrajectories $x^\circ[t_2, T]$, $x^*[t_2, T]$ of which $x^*[t_2, T]$ is optimal with respect to the same criterion $J(\cdot)$ but does not coincide with (nonoptimal) remaining part $x^\circ[t_2, T]$. For totally optimal systems (where the principle of optimality is valid), we have $C[\cdot] \equiv 0$. Otherwise, $C[\cdot]$ should be of higher order than the time difference $t_2 - t_1 \to 0$ for all $t_1 \in [t_o, T)$ in accordance with (7.5) to assure the applicability of dynamic programming methods.

Denote $S = t_2 - t_1$ (same S as in [1, Ch. 9, Section 3,6,13,22]), then (7.12) can be rewritten as follows:

$$C[t_1, t_1 + S] := C\left(u^\circ[t_1, T], u^*[t_1 + S, T]\right) = J\left(t_1, x^\circ[t_1 + S, T]\right) - J\left(x^*[t_1 + S, T]\right), \tag{7.13}$$

where $t_1 = a$ in (7.1) to (7.5), and we get an equivalent representation for (7.5)

$$\lim_{S \to 0} \left\{ \frac{C[t_1, t_1 + S]}{S} \right\} = 0, \qquad \text{for all } t_1 \in [t_o, T). \tag{7.14}$$

Under assumptions of continuous partial derivatives adopted in [1], the functional $C[t_1, t_1 + S]$ is continuously differentiable with respect to S. Consider the limit

$$C(t_1) = \lim_{S \to 0} C[t_1, t_1 + S] = \lim_{S \to 0} \left\{ J\left(t_1, x^\circ[t_1 + S, T]\right) - J\left(x^*[t_1 + S, T]\right) \right\}. \tag{7.15}$$

If $C(t_1) \neq 0$, then in (7.14) we have $\pm\infty$, and noncontiguity follows. If $C(t_1) = 0$, the l'Hôpital rule applies, and (7.14) is equivalent to the condition

$$\lim_{S \to 0} \left\{ \frac{dC[t_1, t_1 + S]}{dS} \right\} = 0, \qquad \text{for all } t_1 \in [t_o, T), \tag{7.16}$$

which is the contiguity condition for general systems and functionals, not necessarily of type (6.1) and (7.1) considered in [1,3].

Note that by construction, $\frac{dC[t_o, t_o + S]}{dS} = 0$ at $S = 0$ for all systems, but this equality at an isolated point $t_1 = t_o$ does not provide the functional equations and PDEs of dynamic programming. Establishing those functional equations requires the contiguity of all semitrajectories, leading to condition (7.14) or (7.16) for all $t_1 > t_o$ as a necessary and sufficient condition for validity of functional equations and PDEs of dynamic programming. For autonomous systems obeying the principle of optimality ($x^*[t_1, T] \subset x^\circ[t_o, T]$ for all $t_1 \geq t_o$), conditions (7.14),(7.16) are redundant, see Remark 7.1.

REMARK 7.2. The weak invariance hypothesis (coincidence in the limit) seems less restrictive than original Bellman's invariance hypothesis (coincidence over the whole time period of the process, see Observation 3.8 in Section 3). This hypothesis for variable $a = t$ in (7.1)–(7.4) may coincide with Bellman's hypothesis, but a proof is needed. This problem is beyond the scope of the paper.

REMARK 7.3. Verification of the contiguity condition (7.16), or (7.14), or (7.5), requires optimization of two different trajectories with variable parameters t_1, t_2. This task may be complicated even for simple systems, see Section 9. For this reason, the contiguity condition is studied here not as a tool in numerical procedures but rather for the understanding of the basis, of the underlying substructure in the theory of dynamic programming.

8. NONCONTIGUOUS OPTIMAL DYNAMICAL SYSTEMS

It is tempting to think that dynamic programming, applicable even without the principle of optimality, might be universally applicable to almost all systems which most likely satisfy the invariance hypothesis and the contiguity condition, with the exception of, maybe, some pathological examples. Unfortunately, this is not the case.

Noncontiguity (i.e., nonsatisfaction of the contiguity condition) in optimal dynamical systems is a rather fine property that invalidates the extension (application) of an optimal control over some part of a trajectory for use over another part of the same trajectory. Strangely as it may seem, noncontiguous systems are common in practice. Contiguity of a system can be destroyed by rather mild constraints usually encountered in engineering and economic models.

Systems with incomplete information and/or structure constraints are generally noncontiguous, even if they are autonomous. To illustrate the existence of such systems, we consider an engineering problem, see [11; 12, pp. 4–7], of optimal stabilization of a stationary linear oscillator (engines, turbines, vibratory machines, etc.) which we study in two distinct settings to demonstrate the fundamental difference in regard to optimality that may occur with modifications in the structure of a controller.

EXAMPLE 8.1. (See [11].) In all textbooks on undergraduate mathematics, mechanics, or control theory, the following linear oscillator equation is considered:

$$\frac{d^2x}{dt^2} + x = u, \qquad t \geq 0, \qquad x(0) = x_o, \qquad \frac{dx(0)}{dt} = v_o. \tag{8.1}$$

Denoting $\frac{dx}{dt} = v$ (velocity), we convert (8.1) into the normal form

$$\frac{dx}{dt} = v, \qquad \frac{dv}{dt} = -x + u, \qquad t \geq 0, \qquad x(0) = x_o, \qquad v(0) = v_o. \tag{8.2}$$

The oscillator (8.1) or (8.2) should be stabilized to minimize the functional

$$f(x_o, v_o) = \min J(u) = \int_0^\infty \left(ax^2 + bv^2 + cu^2\right) dt, \qquad a, b, c = \text{const} > 0. \tag{8.3}$$

The system is autonomous, its optimal trajectories are totally optimal, the principle of optimality applies, thus, denoting $f(x, v) = \min J(u)$ for $t > 0$, the Bellman equation follows from (1.9) with $f_T \equiv 0$ for $T = \infty$ (fixed), cf. [1, Ch. IX, Section 3, equation (12)],

$$f_T(x, v) = \min \left[ax^2 + bv^2 + cu^2 + v\frac{\partial f}{\partial x} + (-x + u)\frac{\partial f}{\partial v}\right] = 0. \tag{8.4}$$

With the control space open, it implies for optimal u°, according to (1.11)

$$2cu^\circ + \frac{\partial f}{\partial v} = 0. \tag{8.5}$$

Substituting optimal control u° from (8.5) into (8.4) (which removes the min operation) and solving the nonlinear PDE, we obtain the minimal cost ($J(u)$ of (8.3) is strictly convex) and the unique optimal feedback

$$f(x,v) = (c+a)^{1/2}(b-2cp)^{1/2}x^2 - 2cpxv + \left(bc - 2c^2p\right)^{1/2} v^2, \tag{8.6}$$

$$u^\circ = px - \left(\frac{b}{c} - 2p\right)^{1/2} v, \qquad p = 1 - \left(1 + \frac{a}{c}\right)^{1/2}. \tag{8.7}$$

We see that feedback (8.7) does not depend on initial conditions in (8.2). It means that every remaining part of an optimal trajectory is itself optimal. Also, for fixed T, equations (1.9),(8.4) do not depend on boundary conditions $x(t_o)$, $x(T)$, $0 \le t_o \le T < \infty$, which means that every intermediate part $x[t_o, T]$ of the optimal trajectory is itself optimal (total optimality) with the same control (8.7) (congruence), in agreement with Theorem 5.1. The optimality is attained in the class of linear stationary feedback controls, though the optimization problem in dynamic programming is formulated in the class of general piecewise continuous controls.

REMARK 8.1. Relationship to multiobjective control. Consider $u^\circ = px + kv$ with fixed $p, k =$ const < 0 as defined in (8.7). Without loss of generality, we can set $c = 1$ in (8.3) to (8.7). Then, simple calculation shows that optimal control (8.7) represents the unique global minimizer not only for functional (8.3) with given $a \ge 0$, $b \ge 0$, $c = 1$, but for any functional (8.3) defined by $c = 1$, $a + b = p^2 + k^2$, which functional will have the global minimum value $f(x,v) = |k|(1-p)x^2 + 2|p|xv + |k|v^2$. It means that multiobjective stabilization problem (8.1)–(8.3) with continuum of such functionals (8.3) is balanced in the sense that unique control (8.7) renders global minima to all those functionals defined by $c = 1$ and nonnegative a, b, $a + b = p^2 + k^2$. In practice, this provides a possibility of choosing a, b to assure reasonable values for different criteria. It is interesting that Bellman's equation (8.4) can be applied to solve such multiobjective control problems with continuum of criteria.

EXAMPLE 8.2. (See [11].) Suppose that the space coordinate $x(t)$ cannot be measured (fluid friction which is common in engineering), and thus, the control must be of the form $u = u(v)$.

If $a = 0$ in (8.3), then $p = 0$, and u° of (8.7) is feasible. However, in (8.3) we have $a > 0$, thus, in (8.7) $p < 0$ and u° is infeasible. In the class of controls $u(v)$ depending on velocity, equation (8.4) does not have a solution, though "approximate" solutions with small $|p|$ do exist, all nonoptimal since they actually correspond to small $a > 0$ in (8.3).

In the spirit of dynamic programming and noting result (8.7), let us look for the optimal feedback in the class of controls

$$u = -2qv, \qquad q > 0, \tag{8.8}$$

which represents a nonsingular Lagrange problem of the calculus of variations with two constraints (8.2) if we eliminate u from (8.3) by the substitution $u = x + v'$ from (8.2), and then substitute (8.8) into (8.2).

For this class of controls, one can also write the Bellman equation by simply substituting (8.8) into (8.4), and then considering $q(\cdot)$ as a control. Denoting $u = -2q(\cdot)v = u^\circ(\cdot)$, one gets the same equations (8.4),(8.5) with solution (8.7), nothing new.

However, the optimal regulator of the form (8.8) with $q =$ const (optimal dampening) does exist. For system (8.2), (8.3), (8.8), the characteristic equation is $r^2 + 2qr + 1 = 0$, so that $\operatorname{Re} r_{1,2} < 0$, integral (8.3) is convergent, and, for $q \ne 1$, it has the value

$$J(x_o, v_o, q) = \left[(a+c)x_o^2 + cv_o^2\right] q + \frac{1}{4}(a+b)\left(x_o^2 + v_o^2\right) q^{-1} + ax_o v_o. \tag{8.9}$$

Solutions are nonoscillating for $q \geq 1$ and oscillating for $0 < q < 1$.

From the equation $\frac{\partial J}{\partial q} = 0$, we find the extremal value

$$q_o^2 = \frac{1}{4}(a+b)\frac{(x_o^2 + v_o^2)}{(a+c)x_o^2 + cv_o^2}, \qquad q_o > 0, \tag{8.10}$$

for which $\frac{\partial^2 J}{\partial q^2} > 0$, yielding the minimum in (8.9), see [11, formulae (5.7)–(5.12)]. Since $J(\cdot)$ of (8.9) is continuously differentiable at $q = 1$, for all values of parameters in (8.9), formulae (8.10),(8.9) are valid also for $q = 1$, covering the whole range of possible friction fluid densities.

Comparing (8.7) and (8.8),(8.10), one can see that

(1) u° of (8.7) does not depend on initial data meaning that u° is optimal not only for any remaining part of an optimal trajectory (Bellman's principle of optimality) but also for all trajectories in the feasible space X. It means that equations (8.6),(8.7) establish a property which may be called *field optimality* in the sense that for system (8.1)–(8.3) there exists one single optimal control (8.7) such that every trajectory starting from any point $x \in X = R^2$ is optimal. Moreover, there is a unique surface (8.6) which renders the optimal value $f(x)$ of functional (8.3) on the trajectory starting from that point x.

(2) $u = -q_o v$ of (8.8),(8.10) does depend on initial data x_o, v_o, meaning that, being optimal for the whole trajectory corresponding to those data, it is not optimal for any remaining part thereof, and hence, Bellman's principle of optimality is invalid in the class of controls (8.8), if $a > 0$ in (8.3).

(3) If $a = 0$ in (8.3), we have $u \equiv u^\circ$, see (8.7), (8.8), (8.10), and thus, for this class of functionals in (8.3), Bellman's principle of optimality is valid in the class of controls (8.8) for the particular system (8.2).

Nonexistence of solutions of the Bellman equation (8.4) for system (8.2), (8.3), (8.8), which equation can be formally written as proposed in [1], means that this autonomous system with incomplete information is noncontiguous. Its noncontiguity invalidates the derivation of Bellman's equations presented in [1]. In contrast, this derivation is valid for contiguous systems even without optimality of remaining semitrajectories because the limit operation as $S \to 0$ removes the necessity of exact coincidence (principle of optimality) of two (maybe, different) semitrajectories starting at the same point $x(t)$, $t > t_o$. Those two semitrajectories need not coincide. They must be touching at $x(t+S)$ as $S \to 0$ to the second or higher order, i.e., lie in a *weak* neighborhood of one another (cf. the curvature in the Weierstrass form of the Euler-Lagrange equations). Then, the Bellman equations are valid and usually have a solution.

REMARK 8.2. There are many wishful interpretations of the principle of optimality, with the same idea, similar to the following: *"To optimize a process, one naturally tries to take optimal decisions at any moment of evolution of the process, meaning that all remaining decisions must be optimal."*

Such common sense formulations conveniently avoid the question of frequency of taking optimal decisions different at every moment of the process, and also ignore all subtle points in the derivation of the functional differential equations and PDEs of dynamic programming. Moreover, the notion of taking momentary decisions, apart from its practical realization, is questionable also theoretically. In Example 8.2, it would mean replacing x_o, v_o by current values $x(t)$, $v(t)$, and in this case, feedback (8.8) would be nonlinear, in contradiction with the requirement to stabilize (8.2) by optimal fluid friction which is linear with $q = \text{const}$ as in (8.10). In this example, momentary decisions do not make sense at all.

Unfortunately, ambiguous "common sense" interpretations are popular and difficult to counter with scientific arguments. If something does not work, or contradictions appear, the problem is simply declared improper for the method. However, smoothly running approximate computations may disguise theoretical failures, and in practice, high price may be exacted for blind application of common sense intuition.

9. VERIFICATION OF NONCONTIGUITY

Simple system (8.1),(8.8) has a general solution, see [11, formulae (5.10) for $q \neq 1$], with which the nonsatisfaction of the principle of optimality is demonstrated in [11, formulae (5.13)–(5.19)] for certain particular solution. However, for application of dynamic programming the principle of optimality is not necessary—the contiguity of the system is required. To evaluate the error committed by assuming $g^{a+S}(.) \cong 0$ in (7.4) for a not totally optimal system, thus, imposing a dynamic programming solution when it is not justified, one has to calculate the noncontiguity limit μ in (7.5), or (7.14), or (7.16). Note that μ is not a loss of optimality on various remaining trajectories whose loss is given by (7.12).

Let us compute the noncontiguity limit $\mu(t, a, b, c, x_o, v_o)$ for the solution (8.2), (8.3), (8.8) with $q = \text{const}$ and $a > 0$. Without loss of generality, we can set $c = 1$, and $v_o = 0$ if starting at an extreme position. To simplify the calculations and to provide for a clear picture while illustrating the notion and the computational procedure, we shall compute the noncontiguity limit for a class of cost functionals defined by the restriction $a + b = 4(a + 1)$, for which $q_o = 1$ in (8.10), $J(x_o, 0, 1) = 2(a + 1)x_o^2$ in (8.9), $u^o = -2v$ in (8.8), and the optimal solution (nonoscillating) of (8.2) is

$$x^o(t) = x_o(1 + t)e^{-t}, \quad v^o(t) = -x_o t e^{-t}, \qquad t \geq 0. \tag{9.1}$$

Now, we shall calculate the noncontiguity limit $\mu(t, a, b, 1, x_o, 0)$, $a+b = 4(a+1)$, $a > 0$. Instead of a, $a + S$ in (7.2)–(7.5), or t_1, t_2 in (7.10)–(7.16), we shall use the notations (variables) t, $t + S$. Then, with initial optimal $q_o = 1$, we have, due to (9.1),

$$\begin{gathered} x^o[t + S, \infty] = x_o(1 + t + S)e^{-(t+S)}, \qquad v^o[t + S, \infty] = -x_o(t + S)e^{-(t+S)}, \\ 0 \leq t < t + S < \infty. \end{gathered} \tag{9.2}$$

At the moment $t + S > 0$, we consider $x^* = x^o(t + S)$, $v^* = v^o(t + S)$ of (9.2) as new initial values for which we calculate the new optimal control (8.8),(8.10) with

$$q^{*2}(t + S) = \frac{2(t + S)^2 + 2(t + S) + 1}{h(t + S)^2 + 2(t + S) + 1}, \qquad q^* > 1, \quad \text{for } t + S > 0, \tag{9.3}$$

where $h = (a + 2)/(a + 1) < 2$ for $a > 0$. Since $q^*(t + S) \to 1$ as $t + S \to 0$, we see that trajectory (9.1) corresponding to $q = 1$ is, indeed, optimal. Since $q^*(t+S) \neq 1$ for any $t+S > 0$, we see that any remaining part of (9.1) is not optimal, and thus, the principle of optimality does not hold on this trajectory. However, dynamic programming equations may still be valid if trajectory (9.1) is contiguous, i.e., satisfies the contiguity condition (7.5), or (7.14), or (7.16). To check its contiguity, we take a point (x^*, v^*) on that trajectory at a moment $t + S > 0$, and, according to the general procedure exhibited in (7.2),(7.3), compute the second integral $J(x^*, v^*, q^*(t))$ of (7.2) corresponding to the original control $u[a, T] \equiv u[t, \infty) = u^o = -2v$ with $q^*(t) = q_o \equiv 1$ for trajectory (9.1), starting at $t + S = a + S$ of (7.2), and its optimal value $J(x^*, v^*, q^*(t + S)) = f^{a+S}(\cdot)$ of (7.3) with new control parameter $q^*(t + S) > 1$ corresponding to optimal control $u^o[t + S, \infty)$, with formula (8.9) which is valid for any $q > 0$ and any initial point (x_o, v_o).

We have

$$J\left(x^o[t+S,\infty]\right) = J\left(x^*, v^*, q^*(t)\right) = x_o^2 e^{-2(t+S)} \Big\{ \left[(t+S)^2 + (a+1)(1+t+S)^2\right] q^*(t) \\ + (a+1)\left[(t+S)^2 + (1+t+S)^2\right] q^{*-1}(t) - a(t+S)(1+t+S) \Big\}, \tag{9.4}$$

$$J\left(x^*[t+S,\infty]\right) = J\left(x^*, v^*, q^*(t+S)\right) = x_o^2 e^{-2(t+S)} \Big\{ \left[(t+S)^2 + (a+1)(1+t+S)^2\right] q^*(t+S) \\ + (a+1)\left[(t+S)^2 + (1+t+S)^2\right] q^{*-1}(t+S) \\ - a(t+S)(1+t+S) \Big\}, \tag{9.5}$$

where $q^*(t) = q_o \equiv 1$ as defined by (8.10) with $c = 1$, $v_o = 0$, $a + b = 4(a+1)$. Now, we can write down the loss of optimality functional (7.13),

$$C[t, t+S] = J\left(t, x^o[t+S,\infty]\right) - J\left(x^*[t+S,\infty]\right) = J\left(x^*, v^*, 1\right) - J\left(x^*, v^*, q^*(t+S)\right) \\ = x_o^2 e^{-2(t+S)} \Big\{ \varphi(t+S)\left[1 - q^*(t+S)\right] + (a+1)\theta(t+S)\left[1 - q^{*-1}(t+S)\right] \Big\}, \tag{9.6}$$

where $\varphi(t+S)$, $\theta(t+S)$ are the corresponding brackets in (9.4),(9.5). Simplification of (9.6) with the notations $q^*(t+S) = Q(t+S)/H(t+S)$, where Q, H are positive square roots of the brackets in (9.3), yields

$$C[t, t+S] = J\left(x^*, v^*, 1\right) - J\left(x^*, v^*, q^*(t+S)\right) = x_o^2 e^{-2(t+S)} \frac{[Q(t+S) - H(t+S)]^2}{(a+1)} \\ = x_o^2 e^{-2(t+S)} \frac{(h-2)^2 (t+S)^4}{[Q(t+S) + H(t+S)]^2 (a+1)} \\ = a^2 x_o^2 e^{-2(t+S)} \frac{(t+S)^4}{[Q(t+S) + H(t+S)]^2 (a+1)^3} > 0, \qquad \text{if } t > 0, \quad S \geq 0, \tag{9.7}$$

which is natural since $J(t, x^o[t+S,\infty])$ produced by the optimal control $u^o = -2v$ is optimal at $t + S = 0$, and nonoptimal at $t + S > 0$. Thus, we have the noncontiguity limit

$$\mu(t, a, b, 1, x_o, 0) = \mu(t, x_o) = \lim_{S \to 0} \left(\frac{C[t, t+S]}{S} \right) = +\infty, \qquad \text{for } t > 0, \tag{9.8}$$

and we see from (9.7) that $\mu(t, x_o) = 0$ as $t = 0$, or in the limit as $t \to \infty$ (which is always the case, by construction, see Section 7), and $\mu(t, x_o) = +\infty$, otherwise, for noncontiguous trajectory (9.1) which renders global minima (of different values) for any functional (8.3) with $c = 1$, $a + b = 4(a+1)$, and restriction (8.8) for controls.

10. APPLICATION OF DYNAMIC PROGRAMMING

There are dozens of books and thousands of papers on dynamic programming. Despite a large number of successful applications, many open problems remain. In this paper, some questions are revisited that have been unanswered for years. Since the theory of dynamic programming is still unfinished, the users should be very careful in its application to practical problems.

Suppose that an optimization or control problem with one objective is considered which is described by a model that includes a multistage process and/or a system of ordinary differential or difference equations with some constraints, boundaries, structural limitations, etc., that may limit free variations of choice parameters, and/or control functions, and/or state trajectories. Then, the following steps can be taken to avoid false solutions created by the application of dynamic programming in a wrong way, or to problems to which it is inapplicable.

STEP 1. Try to adjoin all restrictions and limitations to the system of equations using a sort of Lagrangian, or by transfer of the problem into an appropriate space where variations would be free in the interior of the admissible region. If it is impossible, leave nonadjoined restrictions aside for the moment.

STEP 2. If the so-obtained equivalent system is not autonomous (i.e., explicitly depends on time or on stage number), try to transform or reduce it to a new equivalent problem with autonomous functional and system equations, without introducing additional restrictions, see Remarks 6.1, 6.2. If it is successful, go to Step 5.

STEP 3. If Step 2 is unsuccessful, verify, if possible, the contiguity condition. If it is satisfied, then functional equations and PDEs of dynamic programming are applicable, go to Step 5 (the principle of optimality may be invalid, but it does not matter). If the verification is difficult, go to Step 5.

STEP 4. If the contiguity condition is not satisfied, abandon the use of dynamic programming, and try other methods. End.

STEP 5. Write down the Bellman functional equation (1.9), or (1.16) (often called the Hamilton-Jacobi-Bellman equation), and corresponding PDEs (1.10),(1.11), or (1.17),(1.18). *Remark*: if those PDEs are contradictory, or do not define an optimal control, or other hurdles occur, see Example 6.1, this does *not* mean nonexistence of optimal solutions. In such a case, try other methods, e.g., [6–10,12–14].

STEP 6. If you are sure that the optimal solution, i.e., the optimal control (decisions) and the optimal trajectory (process), are both in the interior of the admissible region, solve those PDEs by available methods. If you are not sure that the optimal solution is in the interior of the admissible regions, write the Karush-Kuhn-Tucker type conditions instead of (1.10),(1.11), or (1.17),(1.18), and solve them by available methods, see Remark 1.4. The solutions obtained are only stationary, not necessarily optimal.

STEP 7. For all those solutions, verify sufficient conditions, e.g., strict local convexity or concavity, see Remark 1.1. If they are satisfied, the problem is solved, go to Step 8. If, in the first approximation, they are not strict but not violated (just $\leq$, or $\geq$), investigate the higher-order terms, and go to Step 8 if sufficient conditions are finally satisfied. If they are violated, then only a stationary solution is found which can be accepted as a suboptimal solution or used as the first approximation for further optimization by other methods, see, e.g., [12]. End of dynamic programming application, go to Step 8.

STEP 8. For a local optimal or stationary solution determined, check the restrictions ignored in Step 1, if any. If at least one of them is not satisfied, then try to reconcile the situation (by reformulating violated constraints or limitations). If it is not possible, either adjust the solution to satisfy all the constraints and limitations (at some loss of optimality in the cost functional), or reconsider the problem using other methods. End.

If your application is described by a model that contains PDEs, or delay equations, or multi-objectives, or some other kind of mathematical objects not yet included in the framework of the classical dynamic programming, try other methods, or wait until further research opens a way for the use of dynamic programming in such applications.

11. CONCLUDING REMARKS

In this paper, the notions of total optimality and contiguity are defined, and a study of contiguity of dynamical systems is presented in application to dynamic programming.

In regard to optimality with a single criterion, dynamical systems can be classified as

(a) totally optimal (e.g., some autonomous systems),
(b) contiguous but not totally optimal,
(c) noncontiguous systems.

It is demonstrated that the principle of optimality is a partial statement about total optimality. With free variations of controls, optimal autonomous systems are usually totally optimal. There are optimal nonautonomous systems that are also totally optimal. Whether or not every such system can be transformed (by a nonautonomous transformation) into a totally optimal system, or reduced to such a system, is an open question of much practical importance.

The dynamic programming approach does not require total optimality: a much weaker contiguity condition would suffice. This condition is necessary and sufficient for validity of dynamic programming. Application of dynamic programming without regard to contiguity means that, in case feasible solutions of the Hamilton-Jacobi-Bellman equation do exist for a noncontiguous system, those solutions are all suboptimal, that is, they are not the best but, maybe, better than some other solutions.

Many practical problems with structural constraints, incomplete information, and other specific requirements do not satisfy the contiguity condition, even if such systems are autonomous. In some cases, optimal problems for such systems can be reformulated to allow for formal setting of dynamic programming equations. Those equations may have solutions which may not respect the constraints. It is interesting to find out how much the costs on real trajectories (with constraints) would exceed the costs of the formal dynamic programming solution without constraints, or with constraints partially satisfied.

Many authors emphasize close connections between the calculus of variations, the max/min principle, and dynamic programming, see [8, p. 91] and also [13–15]. Consequently, the invalidity of dynamic programming for certain classes of optimal problems has implications for the two preceding general methods in control. The noncontiguity affects the validity of certain optimization methods, and must be taken into account in the development and application of optimization and control theories and of variational principles.

REFERENCES

1. R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, NJ, (1957).
2. The Institute, On IEEE people, 15, (July 1984).
3. R.E. Bellman, I. Glicksberg and O.A. Gross, *Some Aspects of the Mathematical Theory of Control Processes*, RAND Corporation, Santa Monica, CA, (1958).
4. R. Bellman, *Adaptive Control Processes: A Guided Tour*, Princeton University Press, Princeton, NJ, (1961).
5. R.E. Bellman and S.E. Dreyfus, *Applied Dynamic Programming*, Princeton University Press, Princeton, NJ, (1962).
6. E. Lee and L. Marcus, *Foundations of Optimal Control Theory*, Wiley, New York, (1967).
7. G. Leitmann, *The Calculus of Variations and Optimal Control*, Plenum Press, New York, (1981).
8. G. Knowles, *An Introduction to Applied Optimal Control*, Academic Press, (1981).
9. A.E. Bryson, Jr. and Y.-C. Ho, *Applied Optimal Control*, Blaisdell, Waltham, MA, (1969).
10. E.A. Galperin, Translational symmetry of certain integrals with application to the calculus of variations and optimal control, *Mathl. Comput. Modelling* **36** (6), 717–728, (2002).
11. E.A. Galperin and E.I. Dergacheva, Optimum stabilization of a linear system with incomplete information, *Automation and Remote Control* **8**, 1193–1202, (1968).
12. E.A. Galperin and Q. Zheng, *Global Solutions in Optimal Control and Games*, NP Research Publ., Montreal, (1991).
13. V.A. Troitskii, On variational problems of optimization of control processes, *Applied Mathematics and Mechanics (PMM)* **26**, (1962).
14. I.N. Roitenberg, *Theorie du Controle Automatique*, Mir, Moscow, (1974).
15. F.H. Clarke and R.B. Vinter, The relation between the maximum principle and dynamic programming, *SIAM Journal on Control and Optimization* **25**, 1291–1311, (1987).